

# Wong-Zakai Approximation through Rough Paths in the G-expectation Framework

Shige Peng

*School of Mathematics, Shandong University,  
Jinan, China, peng@sdu.edu.cn*

Huilin Zhang

*School of Mathematics, Shandong University,  
Jinan, China, huilinzhang2014@gmail.com*

**Abstract:** In this paper, we first build the equivalence between the solution for RDEs driven by lifted G-Brownian motion and that for the corresponding Stratonovich SDE in G-framework by the Wong-Zakai approximation for SDEs in G-framework. Furthermore, the  $\hat{c} - q.s.$  convergence rate with mesh-size  $\frac{1}{n}$  in the  $\alpha$ -Hölder norm is calculated as  $(\frac{1}{n})^\theta$ , where  $\alpha \in (\frac{1}{3}, \frac{1}{2})$ , and  $\theta < \frac{1}{2} - \alpha$ .

**Key words:** G-expectation, rough paths, Wong-Zakai approximation, G-Stratonovich SDEs.

**Mathematics Subject Classification** (2010). 60G46; 60H05; 60G48

## 1 Introduction

G-expectation theory is introduced by Peng in [14], [15]. On the one hand, G-expectation theory could be thought as a helpful tool to handle practical problems concerning probability uncertainty. On the other hand, it also provides a coexistence framework for a set of mutually singular martingale measures. More details about G-expectation theory and the corresponding sublinear expectation theory could be obtained in [16].

Wong and Zakai first built the approximation of Stratonovich SDEs by a sequence of ODEs in [18], [19]. Then Ikeda and Watanabe improved the result in their book [11]. In [12], Lyons established the rough path theory, and used Wong-Zakai theorem to get the equivalence between RDE solutions and SDE solutions. On the other hand, one could also build the equivalence by proving the rough integral and the stochastic integral equal (see [3], [4]). The convergence rate for the Wong-Zakai kind of approximation (not limited to this particular

approximation and Brownian motion) was studied in [6], [2], [5]. However, in G-framework, to get the unified version of solutions under singular martingale measures, the random variable space  $L_G(\Omega_T)$  and process space  $M_G(\Omega_T)$  seem smaller than the classical case (see the next part for the description of these spaces), which makes it less obvious to see whether RDE solutions actually are in this random variable and process spaces. To be more precisely, suppose  $Y_t(\omega)$  solves the following RDE driven by the lifted G-Brownian motion  $\mathbf{B}$ ,

$$dY_t = f(Y) d\mathbf{B}.$$

One fails to directly come to the conclusion that for any  $t \in [0, T]$ ,  $Y_t$  is quasi-continuous (which is a necessary condition for  $Y_t$  to be in  $L_G(\Omega_T)$ , see the next part for the definition of quasi-continuous) only by adaptivity and integrability of RDE solutions, since the quasi-surely continuity is built on uniform topology while rough path theory stands on the p-variation or  $\alpha$ -Hölder topology.

G-Brownian motion was first lifted as a geometric rough path, the limit of smooth paths under p-variation norm, in [10]. In that paper, authors also established the Euler-Maruyama approximation for SDEs. Later, authors of this paper started from the view of Gubinelli [8], [9], and lifted the G-Brownian motion by Kolmogorov's theorem (rough paths version in G-framework) in [17]. Furthermore, G-Stratonovich integral was brought about in [17]. Followed by that work, we establish the equivalence between RDE solutions and G-Stratonovich SDE solutions by applying Wong-Zakai schemes and the universal continuity for rough paths. Firstly, we show that solutions of ODEs driven by piecewise-linear G-Brownian motion actually lie in the random variable space in G-framework. Secondly we use the Wong-Zakai argument to show the convergence of ODEs solutions to Stratonovich SDEs solutions. Thirdly, by applying rough path results, we prove these ODEs solutions also convergence to RDEs solutions in G-framework. Furthermore, the quasi-surely convergence rate of Wong-Zakai approximation is calculated by rough paths techniques.

The paper is organized as the following. In Section 2, we recall some basic notations in G-expectation theory and rough path theory. Then in Section 3, we give the main results of this paper. Firstly we give the description of G-Stratonovich SDE solutions. Secondly, we prove the Wong-Zakai theorem in G-framework. Finally, we calculated the convergence rate by the continuity theorem of the Itô-Lyons mapping.

## 2 Preliminaries about G-expectation and Rough Path

In this part, we give some definitions and conclusions on G-expectation and rough path theories without details and proofs. Readers unfamiliar with these results are recommended to search for details in lecture notes as [3], [13], [14], [16].

## 2.1 The rough path theory

Here rough path theory is presented in the version of Hölder continuity (see Gubinelli [8], Friz and Hairer [3]), rather than the original  $p$ -variation version. Symbols here are adopted from [3].

Furthermore, we refer any Banach space  $V$  as finite dimension space, i.e.  $\mathbb{R}^d$ . The symbol  $V \otimes W$  means the algebraic tensor of two Banach spaces. For any path on some interval  $[0, T]$  with values in a Banach space  $V$ , its  $\alpha$ -Hölder norm is defined as

$$\|X\|_\alpha = \sup_{0 \leq s < t \leq T} \frac{|X_{s,t}|}{|t - s|^\alpha}.$$

Here and from here on,  $X_{s,t} = X_t - X_s$ , for any path  $X$ .

Denote  $\mathcal{C}^\alpha([0, T], V)$  as the space of paths with finite  $\alpha$ -Hölder norm with values in  $V$ . Similarly, a mapping  $\mathbb{X}$  from  $[0, T]^2$  to  $V \otimes V$  is attached with norm

$$\|\mathbb{X}\|_{2\alpha} = \sup_{0 \leq s \neq t \leq T} \frac{|\mathbb{X}_{s,t}|}{|t - s|^{2\alpha}},$$

whenever it's finite.

A rough path on some interval  $[0, T]$  with values in a Banach space  $V$  includes a "rough" continuous path  $X : [0, T] \rightarrow V$ , along with its "iterated integration" part  $\mathbb{X} : [0, T]^2 \rightarrow V \otimes V$ , which satisfies "Chen's identity",

$$\mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = X_{s,u} \otimes X_{u,t}, \quad (1)$$

and Hölder continuity.

**Definition 2.1.** *Fixed  $\alpha \in (\frac{1}{3}, \frac{1}{2}]$  and some Banach space  $V$ , the space of rough paths on  $[0, T]$  is consist of pairs  $(X, \mathbb{X})$  satisfying "Chen's identity" and finite  $\alpha$ -Hölder norm and  $2\alpha$ -Hölder norm respective for  $X$  and  $\mathbb{X}$ . Denote  $\mathcal{C}^\alpha([0, T], V)$  as the rough path space and  $\mathbf{X} := (X, \mathbb{X})$  as an element in this space.*

Though rough path space is not necessarily a linear space, the semi-norm and the distance defined as

$$\begin{aligned} \|\mathbf{X}\|_{\mathcal{C}^\alpha} &:= \|X\|_\alpha + (\|\mathbb{X}\|_{2\alpha})^{\frac{1}{2}} \\ \varrho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}) &:= \|X - \tilde{X}\|_\alpha + \|\mathbb{X} - \tilde{\mathbb{X}}\|_{2\alpha} \end{aligned}$$

will satisfy our need. From here on, we suppose  $\alpha \in (\frac{1}{3}, \frac{1}{2})$  for the need of rough integral with respect to G-Brownian motion.

Now we turn to the rough integral. First we should introduce the integrand, controlled paths. Denote  $\mathcal{L}(V, W)$  the Banach space of bounded linear functional from  $V$  to  $W$ .

**Definition 2.2.** A path  $Y \in \mathcal{C}^\alpha([0, T], \bar{V})$  is said to be controlled by a given path  $X \in \mathcal{C}^\alpha([0, T], V)$ , if there exists  $Y' \in \mathcal{C}^\alpha([0, T], \mathcal{L}(V, \bar{V}))$ , such that the remainder term

$$R_{s,t}^Y := Y_{s,t} - Y'_s X_{s,t},$$

satisfies  $\|R^Y\|_{2\alpha} < \infty$ . Denote the collection of controlled rough paths as  $\mathcal{D}_X^{2\alpha}([0, T], \bar{V})$ . In addition,  $Y'$  is called the Gubinelli derivative of  $Y$ . For  $(Y, Y') \in \mathcal{D}_X^{2\alpha}([0, T], \bar{V})$ , we define its semi-norm as  $\|Y, Y'\|_{X, 2\alpha} := \|Y'\|_\alpha + \|R^Y\|_{2\alpha}$ .

The next theorems taken from [3] for the definition of rough integral, existence and uniqueness for solutions of RDEs are originated from [8]. Also see the p-variation version in [12], [13], [4].

**Theorem 2.3. (Gubinelli)** Suppose  $\mathbf{X} \in \mathcal{C}^\alpha([0, T], V)$ , and  $(Y, Y') \in \mathcal{D}_X^{2\alpha}([0, T], \mathcal{L}(V, W))$ . Then the following compensated Riemann sum converges.

$$\int_0^T Y d\mathbf{X} := \lim_{|\mathcal{P}| \rightarrow 0} \sum_{(s,t) \in \mathcal{P}} (Y_s X_{s,t} + Y'_s \mathbb{X}_{s,t}), \quad (2)$$

where  $\mathcal{P}$  are partitions of  $[0, T]$ , with modulus  $|\mathcal{P}| \rightarrow 0$ . Furthermore, one has the bound

$$\left| \int_s^t Y_r d\mathbf{X}_r - Y_s X_{s,t} - Y'_s \mathbb{X}_{s,t} \right| \leq K(\|X\|_\alpha \|R^Y\|_{2\alpha} + \|\mathbb{X}\|_{2\alpha} \|Y'\|_\alpha) |t - s|^{3\alpha}, \quad (3)$$

with a constant  $K$  depending only on  $\alpha$ .

**Theorem 2.4.** For any  $T > 0$ ,  $\xi \in W$ ,  $f \in \mathcal{C}_b^3(W, \mathcal{L}(V, W))$  and  $\mathbf{X} \in \mathcal{C}^\alpha([0, T], V)$ , there exists a unique solution  $(Y, f(Y)) \in \mathcal{D}_X^{2\alpha}([0, T], W)$  solving the following RDE,

$$Y_t = \xi + \int_0^t f(Y_s) d\mathbf{X}_s, \quad t < T.$$

In the above theorem, for  $\int_0^t f(Y_s) d\mathbf{X}_s$  to hold meaning, one should note that if  $(Y, Y') \in \mathcal{D}_X^{2\alpha}$ ,  $f \in \mathcal{C}_b^2$ , then

$$(f(Y), Df(Y)Y') \in \mathcal{D}_X^{2\alpha}.$$

The continuity result for RDEs could date back to [12], which showed the uniform continuity of solutions for RDEs. The local Lipschitz estimate was studied in [13], [8], [4], [3].

**Theorem 2.5.** Suppose  $f \in \mathcal{C}_b^3(W, \mathcal{L}(V, W))$ , and  $(Y, f(Y))$  is the unique solution to the RDE in the above theorem. Also, suppose  $(\tilde{Y}, f(\tilde{Y}))$  be the RDE solution driven by another rough path  $\tilde{\mathbf{X}}$  with the starting point  $\xi$ . Assume

$$\|\mathbf{X}\|_{\mathcal{C}^\alpha}, \|\tilde{\mathbf{X}}\|_{\mathcal{C}^\alpha} \leq M, \quad \text{for some constant } M.$$

One has the following local Lipschitz estimate on  $[0, T_0]$ ,

$$\|Y - \tilde{Y}\|_\alpha \leq C(|\xi - \tilde{\xi}| + \varrho_\alpha(\mathbf{X}, \tilde{\mathbf{X}})),$$

where  $C$  depends on  $M, \alpha, f$ , and  $T_0$  depends on  $C$ .

From the above theorem, it is not hard to extend the estimate on finite time interval  $[0, T]$ , if  $M$  is the bound on  $[0, T]$ .

## 2.2 The sublinear expectation and the G-expectation theory

In this part, we would give a concrete description of random variables space  $L_G$  and process space  $M_G$ , along with some important properties in G-expectation theorem. Most of these definitions and properties are taken from [16], [1].

Let  $\Omega$  be a given set and  $\mathcal{H}$  be a linear space of real valued functions on  $\Omega$  containing constants. Furthermore, suppose  $\phi(X_1, \dots, X_n) \in \mathcal{H}$  if  $X_1, \dots, X_n \in \mathcal{H}$  and  $\phi \in \mathcal{C}_{l.lip}(\mathbb{R}^n)$ , the space of local Lipschitz functions. The space  $\Omega$  is viewed as sample space and  $\mathcal{H}$  is the space of random variables.

**Definition 2.6.** A sublinear expectation  $\hat{\mathbb{E}}$  is a functional  $\hat{\mathbb{E}} : \mathcal{H} \rightarrow \mathbb{R}$  satisfying:

- $\hat{\mathbb{E}}(c) = c, \quad \forall c \in \mathbb{R};$
- $\hat{\mathbb{E}}(\lambda X) = \lambda \hat{\mathbb{E}}(X), \quad \lambda \geq 0 \quad X \in \mathcal{H};$
- $\hat{\mathbb{E}}(X + Y) \leq \hat{\mathbb{E}}(X) + \hat{\mathbb{E}}(Y), \quad X, Y \in \mathcal{H};$
- $\hat{\mathbb{E}}(X_1) \geq \hat{\mathbb{E}}(X_2) \quad \text{if } X_1 \geq X_2.$

The triple  $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$  is called a sublinear expectation space.

According to the definition of  $\hat{\mathbb{E}}$ , it is simple to check the following lemma.

**Lemma 2.7.** For any  $X \in \mathcal{H}$  with  $\hat{\mathbb{E}}[X] = -\hat{\mathbb{E}}[-X]$ , it holds that,

$$\hat{\mathbb{E}}[X + Y] = \hat{\mathbb{E}}[X] + \hat{\mathbb{E}}[Y], \quad \forall Y \in \mathcal{H}. \quad (4)$$

In particular, if  $\hat{\mathbb{E}}[X] = -\hat{\mathbb{E}}[-X] = 0$ ,

$$\hat{\mathbb{E}}[X + Y] = \hat{\mathbb{E}}[Y], \quad \forall Y \in \mathcal{H}. \quad (5)$$

**Definition 2.8.** In a sublinear expectation space  $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ , a random vector  $Y = (Y_1, \dots, Y_n)$ ,  $Y_i \in \mathcal{H}$ ,  $i = 1 \dots n$ , is said to be independent of another random vector  $X = (X_1, \dots, X_m)$ ,  $X_i \in \mathcal{H}$  under  $\hat{\mathbb{E}}[\cdot]$ , denoted by  $Y \perp X$ , if for every test function  $\varphi \in \mathcal{C}_{b.Lip}(\mathbb{R}^m \times \mathbb{R}^n)$ , we have  $\hat{\mathbb{E}}[\varphi(X, Y)] = \hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(x, Y)]_{x=X}]$ .

**Remark 2.9.** If  $Y \perp X$ , one fails to get the conclusion  $X \perp Y$  automatically. Indeed, this is a main difference between G-expectation theory and the classical case. There are nontrivial examples explaining this point. See Chapter 1 in [16].

**Definition 2.10.** Let  $X_1$  and  $X_2$  be two  $n$ -dimensional random vectors defined respectively in sublinear expectation spaces  $(\Omega_1, \mathcal{H}_1, \hat{\mathbb{E}}_1)$  and  $(\Omega_2, \mathcal{H}_2, \hat{\mathbb{E}}_2)$ . They are called identically distributed, denoted by  $X_1 \stackrel{d}{=} X_2$ , if  $\hat{\mathbb{E}}_1[\varphi(X_1)] = \hat{\mathbb{E}}_2[\varphi(X_2)]$ , for all  $\varphi \in \mathcal{C}_{b.Lip}(\mathbb{R}^n)$ .

**Definition 2.11.** (*G-normal distribution*) A  $d$ -dimensional random vector  $X = (X_1, \dots, X_d)$  in a sublinear expectation space  $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$  is called *G-normally distributed* if  $\hat{\mathbb{E}}[|X|^3] < \infty$  and for each  $a, b \geq 0$

$$aX + b\bar{X} \stackrel{d}{=} \sqrt{a^2 + b^2}X,$$

where  $\bar{X}$  is an independent copy of  $X$ , i.e.,  $\bar{X} \stackrel{d}{=} X$ ,  $\bar{X} \perp X$ , and

$$G(A) := \frac{1}{2} \hat{\mathbb{E}}[\langle AX, X \rangle] : \mathbb{S}_d \rightarrow \mathbb{R},$$

Here  $\mathbb{S}_d$  denotes the collection of  $d \times d$  symmetric matrices.

By Theorem 1.6 in Chapter 3 of [16], we know that if  $X = (X_1, \dots, X_d)$  is G-normal distributed,  $u(t, x) := \hat{\mathbb{E}}[\varphi(x + \sqrt{t}X)]$ ,  $(t, x) \in [0, \infty) \times \mathbb{R}^d$ , with  $\varphi \in \mathcal{C}_{b.Lip}(\mathbb{R}^d)$ , is the viscosity solution of the following G-heat equation:

$$\partial_t u - G(D_x^2 u) = 0, \quad u(0, x) = \varphi(x),$$

with function  $G$  defined as above. Here is an important fact about the distribution of G-normal distributed random variables.

**Lemma 2.12.** Suppose  $X$  is G-normal distributed with  $\hat{\mathbb{E}}[X^2] = \bar{\sigma}^2$  and  $-\hat{\mathbb{E}}[-X^2] = \underline{\sigma}^2$ . Then for each convex function  $\varphi$ , concave function  $\psi$ , one has

$$\hat{\mathbb{E}}[\varphi(X)] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \varphi(\bar{\sigma}x) \exp(-\frac{x^2}{2}) dx; \quad (6)$$

$$\hat{\mathbb{E}}[\psi(X)] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \psi(\underline{\sigma}x) \exp(-\frac{x^2}{2}) dx \quad (7)$$

which means the distribution of G-normal distributed random variable acts as the classical normal distributed one when the distribution function is convex or concave.

Now let  $\Omega = \mathcal{C}^0(\mathbb{R}^+, \mathbb{R}^d)$ , the space of  $\mathbb{R}^d$  valued continuous paths  $(\omega)_{t \geq 0}$  vanishing at the origin. We denote  $B_t$  the canonical process in  $\Omega$ , and  $\hat{\mathbb{E}}$  the G-expectation on the space  $(\Omega, L_{ip}(\Omega_T))$  (see Chapter 3 in [16] for how to define  $\hat{\mathbb{E}}$  through solutions of G-heat equations), where  $L_{ip}(\Omega_T) := \{\phi(B_{t_1 \wedge T}, \dots, B_{t_k \wedge T}) : k \in \mathbb{N}, t_1, \dots, t_k \in [0, \infty), \phi \in \mathcal{C}_{b.Lip}(\mathbb{R}^{k \times d})\}$ , for any  $T > 0$ . Moreover, define  $L_{ip}(\Omega) := \bigcup_{n=1}^{\infty} L_{ip}(\Omega_n)$ . According to the definition of  $\hat{\mathbb{E}}$ , one could define the time consistent conditional expectation  $\hat{\mathbb{E}}[\cdot | \Omega_t]$  as the mapping from  $L_{ip}(\Omega)$  to  $L_{ip}(\Omega_t)$ , for any  $t \geq 0$ . Here is a collection of properties for this mapping.

- $\hat{\mathbb{E}}[\xi | \Omega_t] = \xi$ , for any  $\xi \in L_{ip}(\Omega_t)$ .
- $\hat{\mathbb{E}}[X + Y | \Omega_t] \leq \hat{\mathbb{E}}[X | \Omega_t] + \hat{\mathbb{E}}[Y | \Omega_t]$ .
- $\hat{\mathbb{E}}[\xi X | \Omega_t] = \xi^+ \hat{\mathbb{E}}[X | \Omega_t] + \xi^- \hat{\mathbb{E}}[-X | \Omega_t]$ , for any  $\xi \in L_{ip}(\Omega_t)$

- $\hat{\mathbb{E}}[\hat{\mathbb{E}}[X|\Omega_t]|\Omega_s] = \hat{\mathbb{E}}[X|\Omega_{t \wedge s}]$ , specially,  $\hat{\mathbb{E}}[\hat{\mathbb{E}}[X|\Omega_t]] = \hat{\mathbb{E}}[X]$ .
- $\hat{\mathbb{E}}[X|\Omega_t] = \hat{\mathbb{E}}[X]$ , if  $X$  is independent of  $L_{ip}(\Omega_t)$ .
- $\hat{\mathbb{E}}[X + \xi|\Omega_t] = \hat{\mathbb{E}}[X|\Omega_t] + \xi$ , for any  $\xi \in L_{ip}(\Omega_t)$ ,  $X \in L_{ip}(\Omega)$ .

In the case when  $\hat{\mathbb{E}}[B_1^2] = -\hat{\mathbb{E}}[-B_1^2] = 1$ , the function  $G$  is linear, so  $\hat{\mathbb{E}}$  is linear in this case and G-framework is the classical Wiener case.

For each  $p \geq 1$ ,  $L_G^p(\Omega_T)$  denotes the completion of the linear space  $L_{ip}(\Omega_T)$ , under norm  $\|\cdot\|_{L_G^p} := \{\hat{\mathbb{E}}[|\cdot|^p]\}^{\frac{1}{p}}$ . Obviously, for any  $p \leq q$ ,  $\|\cdot\|_{L_G^p} \leq \|\cdot\|_{L_G^q}$  and  $L_G^q \subseteq L_G^p$ . Furthermore, the conditional expectation  $\hat{\mathbb{E}}[\cdot|\Omega_t]$  could be continuously extended to a mapping from  $L_G^1(\Omega)$  to  $L_G^1(\Omega_t)$  and the extended mapping adopts the above properties.

The following representation theorems for G-expectation  $\hat{\mathbb{E}}$  and  $L_G^1$  is taken from [1].

**Theorem 2.13.** *Assume  $\Gamma$  is a bounded, convex and closed subset of  $\mathbb{R}^{d \times d}$ , which represents function  $G$ , i.e.,*

$$G(A) = \frac{1}{2} \sup_{\gamma \in \Gamma} \text{tr}(A\gamma\gamma'), \text{ for } A \in \mathbb{S}^d.$$

Denote  $P^0$  the Wiener measure and  $B_{s,t} := B_t - B_s$ . Then, for any  $T > 0$  and  $0 = t_0 < t_1 \dots < t_k \leq T$ , the G-expectation has the following representation

$$\begin{aligned} \hat{\mathbb{E}}[\phi(B_{t_0,t_1}, \dots, B_{t_{k-1},t_k})] &= \sup_{a \in \mathcal{A}_{0,T}^\Gamma} E_{P^0}[\phi(\int_0^{t_1} a_s dB_s, \dots, \int_{t_{k-1}}^{t_k} a_s dB_s)] \quad (8) \\ &= \sup_{P^a \in \mathcal{P}_{0,T}^\Gamma} E_{P^a}[\phi(B_{t_0,t_1}, \dots, B_{t_{k-1},t_k})], \quad (9) \end{aligned}$$

where

$$\begin{aligned} \mathcal{P}_{0,T}^\Gamma &:= \{P^a : \text{law of process with the form } \int_0^\cdot a_s dB_s, \text{ for any progressively} \\ &\text{measurable process } \{a_s\}_{s \geq 0}, \text{ with values in } \Gamma, \text{ under Wiener measure } P^0\}. \end{aligned}$$

Furthermore,  $\mathcal{P}_{0,T}^\Gamma$  is tight.

According to this theorem, one could define  $\|\cdot\|_{\mathbb{L}^p} := \sup_{P^a \in \mathcal{P}_{0,T}^\Gamma} E_{P^a}^{\frac{1}{p}}[|\cdot|^p]$ , for any Borel measurable variable, so in the followings, the inequality,

$$\|\xi\|_{\mathbb{L}^p} < \infty,$$

does not mean  $\xi \in L_G^p$ . Actually, there are examples that  $L_G^p$  are nontrivially included in the apparently larger space  $\mathbb{L}^p$ . See [1] for more details.

We define the capacity

$$\hat{c}(A) := \sup_{P \in \mathcal{P}^\Gamma} P(A), \text{ for } A \in \mathcal{B}(\Omega_T).$$

**Definition 2.14.** A property is said to hold "quasi-surely" (q.s.) with respect to  $\hat{c}$ , if it holds true outside a  $\hat{c}$ -polar set (Borel set with capacity 0), and is denoted as  $\hat{c}$ -q.s..

A process  $Y$  on  $[0, T]$  is said to be a quasi-surely modification of another process  $X$  with respect to capacity  $\hat{c}$ , if for any  $t \in [0, T]$

$$Y_t = X_t, \quad \hat{c}\text{-q.s..}$$

If a property stands true  $\hat{c}$ -q.s., then for any  $P \in \mathcal{P}^\Gamma$ , it holds true  $P$ -a.s.. By the definition of  $L_G^p$ , we do not distinguish two random variables if they are equal outside a polar set. The following definitions are important in the characterization of  $L_G^1$ .

**Definition 2.15.** Equip the space  $\Omega_T$  with the uniform topology. A mapping  $X$  on  $\Omega_T$  with values in  $\mathbb{R}$  is said to be quasi-continuous if for any  $\varepsilon > 0$ , there exists an open set  $O$ , with  $\hat{c}(O) < \varepsilon$  such that  $X$  is continuous in  $O^c$ .

**Definition 2.16.** One says that  $X : \Omega_T \rightarrow \mathbb{R}$  has a quasi-continuous version if there exists a quasi-continuous function  $Y$ , such that  $X = Y$   $\hat{c}$ -q.s..

**Theorem 2.17.** One has the following representation for  $L_G^1$ ,

$$L_G^1(\Omega_T) = \{X \in \mathcal{B}(\Omega_T) : X \text{ has a quasi-continuous version, } \lim_{n \rightarrow \infty} \|X|_{\{|X| > n\}}\|_{\mathbb{L}^1} = 0\},$$

where  $\mathcal{B}(\Omega_T)$  is the Borel  $\sigma$  field generated naturally on  $\Omega_T$  by distance. The following proposition is similar as the classical case, so the proof is negligible.

**Proposition 2.18.** Assume that  $(X_n)_{n \geq 1}$  is a sequence in  $L_G^p$ , with  $p > 0$ , and converges to  $X$  in the sense of  $L_G^p$ -norm. Then the convergence holds in the sense of capacity, i.e. for any  $\varepsilon > 0$ ,

$$\hat{c}(|X_n - X| > \varepsilon) \xrightarrow{n} 0.$$

Furthermore, there exists a subsequence  $(X_{n_k})_{k \geq 1}$  converges to  $X$  quasi-surely.

**Remark 2.19.** It is vital to point out that though the above proposition holds true in the  $G$ -framework, even upper-expectation framework, the converse conclusions, i.e. the dominated convergence theorem (the quasi-surely version, in fact, the capacity version still stands true), and the claim that quasi-surely convergence implies convergence in capacity, all fail in  $G$ -framework. More results concerning convergence in  $G$ -framework, relying on the tightness of  $\mathcal{P}^\Gamma$ , could refer to [1].



Here is the stochastic integral (Itô's integral) in G-framework. Fixed  $p \geq 1$ , denote  $M_G^{p,0}(0, T)$  the collection of processes with form

$$\eta_t(\omega) = \sum_{i=0}^{N-1} \xi_i(\omega) 1_{[t_i, t_{i+1})}(t),$$

for a partition  $\{0 = t_0 < \dots < t_N = T\}$  and  $\xi_i \in L_{ip}(\Omega_{t_i}), i = 0 \dots N-1$ . Then define norm  $\|\cdot\|_{M_G^p} := \{\hat{\mathbb{E}} \int_0^T |\eta_s|^p ds\}^{\frac{1}{p}}$  on  $M_G^{p,0}(0, T)$ , and denote  $M_G^p(0, T)$  the completion of  $M_G^{p,0}(0, T)$  under this norm. Consequently, one has the following definition for Itô's integral.

**Definition 2.20.** For each  $\eta \in M_G^{2,0}(0, T)$  with the above form, one has the mapping  $I$  from  $M_G^{2,0}(0, T)$  to  $L_G^2(\Omega_T)$  :

$$I(\eta) = \int_0^T \eta_s dB_s := \sum_{i=0}^{N-1} \xi_i(B_{t_{i+1}} - B_{t_i}) \quad (10)$$

It has been shown (see [14],[15],[16]) that the mapping is continuous and can be extended to the whole space  $M_G^2(0, T)$ , i.e., the mapping  $I$  is a continuous mapping from  $M_G^2(0, T)$  to  $L_G^2(\Omega_T)$ .

Then we define one of the most important processes in G-expectation theory, the quadratic variation processes of G-Brownian motion, denoted as  $\langle B \rangle$ .

$$\langle B \rangle_t := B_t^2 - 2 \int_0^t B_s dB_s. \quad (11)$$

It can be shown that  $\underline{\sigma} \leq \frac{d\langle B \rangle_t}{dt} \leq \bar{\sigma}$ ,  $\hat{c} - q.s.$ , where  $\underline{\sigma} = \sqrt{-\hat{\mathbb{E}}[-B_1^2]}$  and  $\bar{\sigma} = \sqrt{\hat{\mathbb{E}}[B_1^2]}$ . In G-expectation theory,  $\langle B \rangle$  serves the G-Brownian motion as the quadratic variation process, while it shares properties of independent stationary increment just as G-Brownian motion. Moreover, the following integral of a process in  $M_G^{1,0}(0, T)$  can be continuously extended to the completion  $M_G^1(0, T)$ .

$$\int_0^T \eta_t d\langle B \rangle_t := \sum_{i=0}^{N-1} \xi_i(\langle B \rangle_{t_{i+1}} - \langle B \rangle_{t_i}) : M_G^{1,0}(0, T) \rightarrow L_G^1(\Omega_T), \quad (12)$$

where  $\eta$  is defined as above, only  $L_G^2$  replaced by  $L_G^1$ .

For the multi-dimensional case, one could obtain similar results. Indeed, let  $(B_t)_{t \geq 0}$  be a d-dimensional G-Brownian motion. For any  $a \in \mathbb{R}^d$ ,  $B^a := a \cdot B$  is still a  $G_a$ -Brownian motion. Then according to results in one-dimensional case, one could define integrals with respect to  $B^a, \langle B^a \rangle$ , and obtain continuity for these mappings. Furthermore, the mutual variation process  $\langle B^a, B^{\bar{a}} \rangle_t$  could be defined by polarization.

At last, we end this subsection with the famous Itô's formula in G-framework. The proof could also be obtained in [16].

**Theorem 2.21.** *Let  $\Phi$  be a twice continuous differentiable function on  $\mathbb{R}^n$  with polynomial growth for the first and second order derivatives.  $X$  is a Itô process, i.e.*

$$X_t^\nu = X_0^\nu + \int_0^t \alpha_s^\nu ds + \int_0^t \eta_s^{\nu ij} d\langle B^i, B^j \rangle_+ + \int_0^t \beta_s^{\nu j} dB_s^j$$

where  $\nu = 1, \dots, n$ ,  $i, j = 1, \dots, d$ ,  $\alpha_s^\nu, \eta_s^{\nu ij}, \beta_s^{\nu j}$  are bounded processes in  $M_G^2(0, T)$ . Here repeated indices means summation over the same ones. Then for each  $t \geq s \geq 0$  we have in  $L_G^2(\Omega_t)$ :

$$\begin{aligned} \Phi(X_t) - \Phi(X_s) &= \int_s^t \partial_{x^\nu} \Phi(X_u) \beta_u^{\nu j} dB_u^j + \int_s^t \partial_{x^\nu} \Phi(X_u) \alpha_u^\nu du \\ &\quad + \int_s^t [\partial_{x^\nu} \Phi(X_u) \eta_u^{\nu ij} + \frac{1}{2} \partial_{x^\mu x^\nu}^2 \Phi(X_u) \beta_u^{\mu i} \beta_u^{\nu j}] d\langle B^i, B^j \rangle_u. \end{aligned}$$

### 3 Main Result

#### 3.1 G-Stratonovich SDE

The G-Stratonovich integral is introduced in [17] and is proved to be equal with rough integral under certain assumptions. We only consider one dimensional case for simplicity.

Define  $\langle Y, B \rangle_t := \lim_{|\mathcal{P}| \rightarrow 0} \sum_{(u,v) \in \mathcal{P}} Y_{u,v} B_{u,v}$  whenever the limit exists in  $L_G^1(\Omega_t)$ , for any  $t \in [0, T]$ . Then one has the following lemma and corollary concerning the existence of  $\langle Y, B \rangle$ , the proof of which could be obtained in [17].

**Lemma 3.1.** *For  $\beta \in M_G^2$ , assume that there exists a sequence  $(\beta^n)_{n=1}^\infty \in M_G^{2,0}$ , such that  $\sup_{|v-u| \leq \frac{1}{m}, [u,v] \subseteq [0,T]} \hat{\mathbb{E}}[\sup_{t \in [u,v]} |\beta_t^n - \beta_t|^2] \xrightarrow{n,m} 0$ . Then, for  $Y_t := \int_0^t \beta_r dB_r$ , one has*

$$\langle Y, B \rangle_t = \int_0^t \beta_r d\langle B \rangle_r, \quad \hat{c} - q.s.. \quad (13)$$

**Corollary 3.2.** *For*

$$Y_t = \xi + \int_0^t \beta_s dB_s + \int_0^t f_s ds + \int_0^t \gamma_s d\langle B \rangle_s, \quad (14)$$

with  $\beta$  satisfying the condition in the above lemma, and  $f, \gamma \in M_G^{2+\delta}(0, T)$ , for some  $\delta > 0$ , one has the expression,

$$\langle Y, B \rangle_t = \int_0^t \beta_s d\langle B \rangle_s, \quad \hat{c} - q.s..$$

**Definition 3.3.** (*G-Stratonovich integration*) Suppose  $Y \in M_G^2(0, T)$ , and  $\langle Y, B \rangle$  exist. The G-Stratonovich integral of  $Y$  against  $B$ , with value in  $L_G^1$ , is given by identity:

$$\int_0^t Y_s \circ dB_s := \int_0^t Y_s dB_s + \frac{1}{2} \langle Y, B \rangle_t, \quad \hat{c} - q.s.. \quad (15)$$

Now we consider the following G-Stratonovich SDE,

$$X_t = x_0 + \int_0^t f(X_s) \circ dB_s + \int_0^t g(X_s) d\langle B \rangle_s + \int_0^t h(X_s) ds. \quad (16)$$

where  $x_0 \in \mathbb{R}$ , and  $t \in [0, T]$ .

**Theorem 3.4.** Assume  $f, g, h \in C_b^2(\mathbb{R})$ , bounded functions with bounded first and second derivatives. There exists a unique  $X \in M_G^2(0, T)$ , such that  $X$  solves (16) and satisfies

$$\langle f(X), B \rangle_t = \int_0^t f'(X_s) f(X_s) d\langle B \rangle_s.$$

*Proof.* Consider the G-Itô version of (16),

$$\tilde{X}_t = x_0 + \int_0^t f(\tilde{X}_s) dB_s + \int_0^t (g(\tilde{X}_s) + \frac{1}{2} f'(\tilde{X}_s) f(\tilde{X}_s)) d\langle B \rangle_s + \int_0^t h(\tilde{X}_s) ds. \quad (17)$$

According to Theorem 1.2 in Chapter V of [16], there exists a unique  $\tilde{X} \in M_G^2$  solving (17). By the definition of G-Stratonovich integral, to complete the proof, one only needs to show that

$$\langle f(\tilde{X}), B \rangle_t = \int_0^t f'(\tilde{X}_s) f(\tilde{X}_s) d\langle B \rangle_s.$$

By Itô's formula, one obtains

$$\begin{aligned} f(\tilde{X}_t) &= f(x_0) + \int_0^t f'(\tilde{X}_s) f(\tilde{X}_s) dB_s + \int_0^t [f'(\tilde{X}_s)(g(\tilde{X}_s) \\ &\quad + \frac{1}{2} f'(\tilde{X}_s) f(\tilde{X}_s)) + \frac{1}{2} f''(\tilde{X}_s) f^2(\tilde{X}_s)] d\langle B \rangle_s + \int_0^t f'(\tilde{X}_s) h(\tilde{X}_s) ds. \end{aligned}$$

Denote  $\tilde{f}(x) = f'(x) f(x)$ ,  $\beta_s = \tilde{f}(\tilde{X}_s)$ ,  $\beta_s^n = \sum_{i=0}^{k_n-1} \tilde{f}(\tilde{X}_{t_i^n}) 1_{[t_i^n, t_{i+1}^n)}(s)$ , where  $\mathcal{Q}^n = \{0 = t_0 < t_1^n < \dots < t_{k_n}^n = T\}$ , is any sequence of partitions with  $|\mathcal{Q}^n| \rightarrow 0$ . It is clear that  $\tilde{f} \in \mathcal{C}_b^1$ . Then one has the following inequality,

$$\begin{aligned} \sup_{|v-u| \leq \frac{1}{m}} \hat{\mathbb{E}}[sup_{t \in [u, v]} |\beta_t^n - \beta_t|^2] &\leq \sup_{|v-u| \leq \frac{1}{m}} \sum_{\mathcal{Q}^n \cap [u, v]} \hat{\mathbb{E}}[\sup_{s \in [u, v] \cap [t_i^n, t_{i+1}^n)} |\tilde{f}(\tilde{X}_s) - \tilde{f}(\tilde{X}_{t_i^n})|^2] \\ &\leq C \sup_{|v-u| \leq \frac{1}{m}} \sum_{\mathcal{Q}^n \cap [u, v]} \hat{\mathbb{E}}[\sup_{s \in [u, v] \cap [t_i^n, t_{i+1}^n)} |\tilde{X}_s - \tilde{X}_{t_i^n}|^2] \\ &= C \sup_{|v-u| \leq \frac{1}{m}} \sum_{\mathcal{Q}^n \cap [u, v]} \sup_{P^a \in \mathcal{P}_{0, T}^a} E_{P^a}[\sup_{s \in [u, v] \cap [t_i^n, t_{i+1}^n)} |\tilde{X}_s - \tilde{X}_{t_i^n}|^2] \\ &\leq C(\frac{1}{m} + 2|\mathcal{Q}^n|), \end{aligned}$$

the last inequality of which follows from the fact that  $\tilde{X}$  solves (17) and basic martingale inequalities.

By applying Corollary 3.2, one obtains the desired result.  $\square$

**Remark 3.5.** *One could consider equations with driven functions less regular than  $\mathcal{C}_b^2$ . We add such assumptions to get the convergence rate by rough path methods in the next part.*

### 3.2 The Wong-Zakai Approximation in G-framework

In this part, we consider the Wong-Zakai approximation in G-framework. Suppose  $T = 1$ ,  $\{t_j^n\}_{j=0}^n$  is the partition with mesh size  $1/n$  and  $B_t^{(n)}$  is the piecewise linearization of G-Brownian motion  $B_t$  segmented by  $\{t_j^n\}_{j=0}^n$ . Consider the following ODEs with initial condition  $x_0$ .

$$dY_t^{(n)} = f(Y_t^{(n)})dB_t^{(n)} + g(Y_t^{(n)})d\langle B \rangle_t + h(Y_t^{(n)})dt. \quad (18)$$

It is clear that for any  $n$ ,  $Y^{(n)}$  can be defined pathwisely  $\hat{c}$ -quasi surely. We need the first lemma to show  $Y_t^{(n)} \in L_G^1(\Omega_{t_{j+1}^{(n)}})$ ,  $t \in [t_j^{(n)}, t_{j+1}^{(n)}]$ .

**Lemma 3.6.** *Assume  $f, g, h$  bounded and Lipschitz. For fixed  $n$ , one has  $Y_t^{(n)} \in L_G^2(\Omega_{t_{j+1}^{(n)}})$ , for any  $t \in [t_j^{(n)}, t_{j+1}^{(n)}]$ .*

*Proof.* By induction, one only needs to show for any  $n \geq 1$ ,  $j = 0 \dots n-1$ ,  $Y_{t_{j+1}^{(n)}}^{(n)} \in L_G^1(\Omega_{t_{j+1}^{(n)}})$ . In the following proof, we omit  $(n)$  in  $Y^{(n)}$  and  $t_j^{(n)}$  for simplicity, i.e. suppose  $Y_t$  solves

$$Y_t = Y_{t_j} + \frac{B_{t_j, t_{j+1}}}{t_{j+1} - t_j} \int_{t_j}^t f(Y_s) ds + \int_{t_j}^t g(Y_s) d\langle B \rangle_s + \int_{t_j}^t h(Y_s) ds, \quad t \in [t_j, t_{j+1}], \quad (19)$$

pathwisely and  $Y_{t_j} \in L_G^1(\Omega_{t_j})$ .

Consider the discretization of (19),

$$y_t^{(m)} = y_{k-1}^{(m)} + \frac{B_{t_j, t_{j+1}}}{t_{j+1} - t_j} f(y_{k-1}^{(m)})(t - \tau_{k-1}^{(m)}) + g(y_{k-1}^{(m)})\langle B \rangle_{\tau_{k-1}^{(m)}, t} + h(y_{k-1}^{(m)})(t - \tau_{k-1}^{(m)}), \quad t \in [\tau_{k-1}^{(m)}, \tau_k^{(m)}], \quad (20)$$

where  $\{\tau_k^{(m)}\}_{k=0}^m$  is the partition of  $[t_j, t_{j+1}]$  with mesh-size  $\frac{1}{m}$ , and  $y_k^{(m)} = y_{\tau_k^{(m)}}^{(m)}$ .

It holds that

$$\begin{aligned}
Y_{t_{j+1}} - y_{t_{j+1}}^{(m)} &= \sum_{i=0}^{m-1} \left[ \frac{B_{t_j, t_{j+1}}}{t_{j+1} - t_j} \int_{\tau_i^{(m)}}^{\tau_{i+1}^{(m)}} (f(Y_s) - f(y_i^{(m)})) ds \right. \\
&\quad \left. + \int_{\tau_i^{(m)}}^{\tau_{i+1}^{(m)}} (g(Y_s) - g(y_i^{(m)})) d\langle B \rangle_s + \int_{\tau_i^{(m)}}^{\tau_{i+1}^{(m)}} (h(Y_s) - h(y_i^{(m)})) ds \right] \\
&= \sum_{i=0}^{m-1} \left\{ \frac{B_{t_j, t_{j+1}}}{t_{j+1} - t_j} \int_{\tau_i^{(m)}}^{\tau_{i+1}^{(m)}} (f(Y_s) - f(y_s^{(m)})) ds + \int_{\tau_i^{(m)}}^{\tau_{i+1}^{(m)}} (g(Y_s) - g(y_s^{(m)})) d\langle B \rangle_s \right. \\
&\quad \left. + \int_{\tau_i^{(m)}}^{\tau_{i+1}^{(m)}} (h(Y_s) - h(y_s^{(m)})) ds + \frac{B_{t_j, t_{j+1}}}{t_{j+1} - t_j} \int_{\tau_i^{(m)}}^{\tau_{i+1}^{(m)}} (f(y_s^{(m)}) - f(y_i^{(m)})) ds \right. \\
&\quad \left. + \int_{\tau_i^{(m)}}^{\tau_{i+1}^{(m)}} (g(y_s^{(m)}) - g(y_i^{(m)})) d\langle B \rangle_s + \int_{\tau_i^{(m)}}^{\tau_{i+1}^{(m)}} (h(y_s^{(m)}) - h(y_i^{(m)})) ds \right\}.
\end{aligned}$$

Note that by boundedness of  $f, g, h$ , and  $\frac{d\langle B \rangle_t}{dt} \leq \bar{\sigma}, \hat{c} - q.s.$ ,

$$\int_{\tau_i^{(m)}}^{\tau_{i+1}^{(m)}} |y_s^{(m)} - y_i^{(m)}| ds \leq C(1 + \frac{|B_{t_j, t_{j+1}}|}{t_{j+1} - t_j})(\tau_{i+1}^{(m)} - \tau_i^{(m)})^2 \quad (21)$$

where  $C$  depends on  $f, g, h$  and  $\bar{\sigma}$ .

According to (21) and Lipschitzness for  $f, g, h$ , one can simply get that

$$|Y_{t_{j+1}} - y_{t_{j+1}}^{(m)}| \leq C(1 + \frac{|B_{t_j, t_{j+1}}|}{t_{j+1} - t_j}) \int_{t_j}^{t_{j+1}} |Y_s - y_s^{(m)}| ds + C(1 + \frac{|B_{t_j, t_{j+1}}|}{t_{j+1} - t_j}) \frac{1}{m} (t_{j+1} - t_j),$$

where  $C$  depends on  $f, g, h$  and  $\bar{\sigma}$ .

By Gronwall's inequality, one has the following inequality,

$$|Y_{t_{j+1}} - y_{t_{j+1}}^{(m)}| \leq \frac{C}{m} (|t_{j+1} - t_j| + |B_{t_j, t_{j+1}}|) e^{C(t_{j+1} - t_j)} e^{C|B_{t_j, t_{j+1}}|}.$$

Then, one takes the expectation and apply Hölder's inequality, and gets the following result,

$$\hat{\mathbb{E}}|Y_{t_{j+1}} - y_{t_{j+1}}^{(m)}| \leq \frac{C}{m},$$

where  $C$  depends on  $|t_{j+1} - t_j|$ , and increasingly depends on  $\bar{\sigma}$ , the bound for  $f, g, h$ , and their Lipschitz constants.

It is clear that  $y_t^{(m)} \in L_G^1(\Omega_{t_{j+1}}), t \in [t_j, t_{j+1}]$ . This fact and the above convergence imply  $Y_{t_{j+1}} \in L_G^1(\Omega_{t_{j+1}})$ . The conclusion follows by the square integrability of  $Y_{t_{j+1}}$ .  $\square$

Now we prove the Wong-Zakai approximation in G-framework.

**Theorem 3.7.** Suppose  $Y_t^{(n)}$  solves ODEs (18)  $\hat{c}$ -quasi surely, and  $X_t$  solves (16) as defined in Theorem 3.4, with  $f, g, h \in \mathcal{C}_b^2$ . Then for any  $t \in [0, 1]$ ,  $Y_t^{(n)}$  converges to  $X_t$  in  $L_G^2$ -norm sense. Furthermore, for any  $t \in [0, 1]$ , one has the following inequality,

$$\hat{\mathbb{E}}[(Y_t^{(n)} - X_t)^2] \leq K \frac{1}{\sqrt{n}}, \quad (22)$$

where  $K$  depends on  $f, g, h$  and  $\bar{\sigma}$ .

*Proof.* Without loss of generality, we suppose  $g, h = 0$ , for simplicity. Indeed, the argument would be the same but the process would be more lengthy. In the following, the constant  $K$  may be different from line to line. Consider the Maruyama's approximation to G-Stratonovich SDE (16),

$$dX_t^{(n)} = X_{j-1}^{(n)} + f(X_{j-1}^{(n)})B_{t_{j-1}^{(n)}, t} + \frac{1}{2}f'(X_{j-1}^{(n)})f(X_{j-1}^{(n)})\langle B \rangle_{t_{j-1}^{(n)}, t}, \quad t \in [t_{j-1}^{(n)}, t_j^{(n)}], \quad (23)$$

where  $X_{j-1}^{(n)} = X_{t_{j-1}^{(n)}}^{(n)}$ ,  $j = 1 \dots n$ .

By Maruyama's approximation in G-framework, i.e. Theorem 7 in Part 3 of [10], one only needs to show

$$\hat{\mathbb{E}}(Y_t^{(n)} - X_t^{(n)})^2 \leq K \frac{1}{\sqrt{n}}, \quad \forall t \in [0, 1]. \quad (24)$$

Note that for any  $t \in [t_j^{(n)}, t_{j+1}^{(n)})$ , one can simply get the identities

$$\begin{aligned} \hat{\mathbb{E}}(Y_t^{(n)} - Y_j^{(n)})^2 &\leq \frac{K}{n} \\ \hat{\mathbb{E}}(X_t^{(n)} - X_j^{(n)})^2 &\leq \frac{K}{n}. \end{aligned}$$

We only need to prove (24) for  $t = t_j^{(n)}$ . Firstly, one has the following identity,

$$Y_j^{(n)} - X_j^{(n)} = (Y_{j-1}^{(n)} - X_{j-1}^{(n)}) + B_{t_{j-1}^{(n)}, t_j^{(n)}}(f(Y_{j-1}^{(n)}) - f(X_{j-1}^{(n)})) \quad (25)$$

$$+ \frac{B_{t_{j-1}^{(n)}, t_j^{(n)}}}{t_j^{(n)} - t_{j-1}^{(n)}} \int_{t_{j-1}^{(n)}}^{t_j^{(n)}} (f(Y_s^{(n)}) - f(Y_{j-1}^{(n)})) ds \quad (26)$$

$$- \frac{1}{2}f'(X_{j-1}^{(n)})f(X_{j-1}^{(n)})\langle B \rangle_{t_{j-1}^{(n)}, t_j^{(n)}}. \quad (27)$$

By Taylor's expansion, for (26), one has

$$\begin{aligned}
& \frac{B_{t_{j-1}^{(n)}, t_j^{(n)}}}{t_j^{(n)} - t_{j-1}^{(n)}} \int_{t_{j-1}^{(n)}}^{t_j^{(n)}} (f(Y_s^{(n)}) - f(Y_{j-1}^{(n)})) ds \\
&= \frac{B_{t_{j-1}^{(n)}, t_j^{(n)}}}{t_j^{(n)} - t_{j-1}^{(n)}} \int_{t_{j-1}^{(n)}}^{t_j^{(n)}} f'(Y_{\tau_s}^{(n)}) (Y_s^{(n)} - Y_{j-1}^{(n)}) ds \\
&= \left( \frac{B_{t_{j-1}^{(n)}, t_j^{(n)}}}{t_j^{(n)} - t_{j-1}^{(n)}} \right)^2 \int_{t_{j-1}^{(n)}}^{t_j^{(n)}} f'(Y_{\tau_s}^{(n)}) \int_{t_{j-1}^{(n)}}^s f(Y_r^{(n)}) dr ds, \tag{28}
\end{aligned}$$

where  $Y_{\tau_s}^{(n)} = Y_{j-1}^{(n)} + \theta(Y_s^{(n)} - Y_{j-1}^{(n)})$ ,  $\theta \in (0, 1)$ .

By subtracting (27) from (28) and inserting terms, one could obtains

$$\begin{aligned}
& \left( \frac{B_{t_{j-1}^{(n)}, t_j^{(n)}}}{t_j^{(n)} - t_{j-1}^{(n)}} \right)^2 \int_{t_{j-1}^{(n)}}^{t_j^{(n)}} \int_{t_{j-1}^{(n)}}^s [f'(Y_{\tau_s}^{(n)}) f(Y_r^{(n)}) - f'(Y_{\tau_s}^{(n)}) f(Y_{\tau_s}^{(n)})] dr ds \\
&+ \left( \frac{B_{t_{j-1}^{(n)}, t_j^{(n)}}}{t_j^{(n)} - t_{j-1}^{(n)}} \right)^2 \int_{t_{j-1}^{(n)}}^{t_j^{(n)}} \int_{t_{j-1}^{(n)}}^s [f'(Y_{\tau_s}^{(n)}) f(Y_{\tau_s}^{(n)}) - f'(Y_{j-1}^{(n)}) f(Y_{j-1}^{(n)})] dr ds \\
&+ \left( \frac{B_{t_{j-1}^{(n)}, t_j^{(n)}}}{t_j^{(n)} - t_{j-1}^{(n)}} \right)^2 \frac{(t_j^{(n)} - t_{j-1}^{(n)})^2}{2} f'(Y_{j-1}^{(n)}) f(Y_{j-1}^{(n)}) - \frac{1}{2} \langle B \rangle_{t_{j-1}^{(n)}, t_j^{(n)}} f'(Y_{j-1}^{(n)}) f(Y_{j-1}^{(n)}) \\
&+ \frac{1}{2} \langle B \rangle_{t_{j-1}^{(n)}, t_j^{(n)}} f'(Y_{j-1}^{(n)}) f(Y_{j-1}^{(n)}) - \frac{1}{2} f'(X_{j-1}^{(n)}) f(X_{j-1}^{(n)}) \langle B \rangle_{t_{j-1}^{(n)}, t_j^{(n)}}.
\end{aligned}$$

In conclusion, we already get the following identity

$$\begin{aligned}
& Y_j^{(n)} - X_j^{(n)} = (Y_{j-1}^{(n)} - X_{j-1}^{(n)}) + B_{t_{j-1}^{(n)}, t_j^{(n)}} (f(Y_{j-1}^{(n)}) - f(X_{j-1}^{(n)})) \\
&+ \frac{1}{2} \langle B \rangle_{t_{j-1}^{(n)}, t_j^{(n)}} [f'(Y_{j-1}^{(n)}) f(Y_{j-1}^{(n)}) - f'(X_{j-1}^{(n)}) f(X_{j-1}^{(n)})] + \frac{1}{2} f'(Y_{j-1}^{(n)}) f(Y_{j-1}^{(n)}) (B_{t_{j-1}^{(n)}, t_j^{(n)}}^2 - \langle B \rangle_{t_{j-1}^{(n)}, t_j^{(n)}}) \\
&+ \left( \frac{B_{t_{j-1}^{(n)}, t_j^{(n)}}}{t_j^{(n)} - t_{j-1}^{(n)}} \right)^2 \int_{t_{j-1}^{(n)}}^{t_j^{(n)}} (s - t_{j-1}^{(n)}) [f'(Y_{\tau_s}^{(n)}) f(Y_{\tau_s}^{(n)}) - f'(Y_{j-1}^{(n)}) f(Y_{j-1}^{(n)})] ds \\
&+ \left( \frac{B_{t_{j-1}^{(n)}, t_j^{(n)}}}{t_j^{(n)} - t_{j-1}^{(n)}} \right)^2 \int_{t_{j-1}^{(n)}}^{t_j^{(n)}} \int_{t_{j-1}^{(n)}}^s f'(Y_{\tau_s}^{(n)}) [f(Y_r^{(n)}) - f(Y_{\tau_s}^{(n)})] dr ds.
\end{aligned}$$

Denote the above six terms as  $\varepsilon_l^{(n,j)}$ ,  $l = 1 \dots 6$ . Firstly, by  $f, g, h \in \mathcal{C}_b^2$ , it is clear that

$$\begin{aligned}
\hat{\mathbb{E}}(\varepsilon_l^{(n,j)})^2 &\leq K \left( \frac{1}{n} \right)^3, \quad l = 5, 6, \\
\hat{\mathbb{E}}(\varepsilon_l^{(n,j)})^2 &\leq K \left( \frac{1}{n} \right)^2, \quad l = 4.
\end{aligned}$$

Secondly, by Lipschitzness of  $f, g, h$ , one has

$$\hat{\mathbb{E}}(\varepsilon_1^{(n,j)})^2 + \hat{\mathbb{E}}(\varepsilon_2^{(n,j)})^2 + \hat{\mathbb{E}}(\varepsilon_3^{(n,j)})^2 + 2\hat{\mathbb{E}}(\varepsilon_1^{(n,j)}\varepsilon_3^{(n,j)}) \leq (1 + \frac{K}{n})\hat{\mathbb{E}}(Y_{j-1}^{(n)} - X_{j-1}^{(n)})^2.$$

By the above lemma,  $Y_{k-1}^{(n)} \in L_G^2(\Omega_{t_{k-1}^{(n)}})$ , which is independent from  $B_{t_{j-1}^{(n)}, t_j^{(n)}}$  and  $\langle B \rangle_{t_{j-1}^{(n)}, t_j^{(n)}}$ , so one gets

$$\hat{\mathbb{E}}(\varepsilon_1^{(n,j)}\varepsilon_l^{(n,j)}) = 0, \quad l = 2, 4.$$

Also, note that

$$\begin{aligned} \hat{\mathbb{E}}B_{t_{j-1}^{(n)}, t_j^{(n)}}[(B_{t_{j-1}^{(n)}, t_j^{(n)}})^2 - \langle B \rangle_{t_{j-1}^{(n)}, t_j^{(n)}}] &= \hat{\mathbb{E}}[B_{t_{j-1}^{(n)}, t_j^{(n)}} \int_{t_{j-1}^{(n)}}^{t_j^{(n)}} B_{t_{j-1}^{(n)}, r} dB_r] \leq K(\frac{1}{n})^{\frac{3}{2}}, \\ \hat{\mathbb{E}}\langle B \rangle_{t_{j-1}^{(n)}, t_j^{(n)}}[(B_{t_{j-1}^{(n)}, t_j^{(n)}})^2 - \langle B \rangle_{t_{j-1}^{(n)}, t_j^{(n)}}] &\leq K(\frac{1}{n})^2, \end{aligned}$$

and one could obtain

$$\hat{\mathbb{E}}[\varepsilon_4^{(n,j)}\varepsilon_l^{(n,j)}] \leq K(\frac{1}{n})^{\frac{3}{2}}, \quad l = 2, 3.$$

As for other intersection terms concerning  $\varepsilon_5^{(n,j)}, \varepsilon_6^{(n,j)}$ , one can apply Hölder's inequality directly by noticing  $\hat{\mathbb{E}}(\varepsilon_1^{(n,j)})^2$  bounded and obtain

$$\hat{\mathbb{E}}[\varepsilon_k^{(n,j)}\varepsilon_l^{(n,j)}] \leq K(\frac{1}{n})^{\frac{3}{2}}, \quad l = 1 \dots 6, k = 5, 6.$$

Finally, one gets

$$\begin{aligned} \hat{\mathbb{E}}[Y_j^{(n)} - X_j^{(n)}]^2 &\leq (1 + \frac{K}{n})\hat{\mathbb{E}}[Y_{j-1}^{(n)} - X_{j-1}^{(n)}]^2 + K(\frac{1}{n})^{\frac{3}{2}} \quad (29) \\ &\leq \sum_{i=0}^{j-1} (1 + \frac{K}{n})^i K(\frac{1}{n})^{\frac{3}{2}} \leq \frac{K}{\sqrt{n}} \end{aligned}$$

□

**Remark 3.8.** For equations with non-zero  $g, h$ , according to the above argument,  $g, h$  would only increase  $K$  on the right hand of inequality (29).

### 3.3 The Quasi-surely Convergence Rate for Wong-Zakai Approximation

In this part, we will calculate the quasi-surely convergence rate for the Wong-Zakai approximation by rough path theory. According to Theorem 2.5, one needs to calculate  $\varrho(\mathbf{B}, \mathbf{B}^{(n)})$ . The convergence of  $\mathbf{B}^{(n)}$  to  $\mathbf{B}^{strat}$  in the sense of capacity was firstly proved in [10] under the p-variation distance by confining the partition as dyadic. In the following, we resort to Kolmogorov criterion for rough path distance and then we no longer need the partition to be dyadic. Here is the Kolmogorov criterion in G-framework (see Theorem 3.1 and Theorem 3.3 in [3] for the classical case).



**Theorem 3.9.** For fixed  $q \geq 2, \beta > \frac{1}{q}$ , assume  $X(\omega) : [0, T] \rightarrow \mathbb{R}$  and  $\mathbb{X}(\omega) : [0, T]^2 \rightarrow \mathbb{R}$  are processes with  $X_t \in L_G^q(\Omega_T), \mathbb{X}_{s,t} \in L_G^{\frac{q}{2}}(\Omega_T), \forall s, t \in [0, T]$ , and satisfy relation (1) quasi-surely. If for any  $s, t \in [0, T]$ , one has bounds

$$\|X_{s,t}\|_{L_G^q} \leq C|t-s|^\beta, \quad \|\mathbb{X}_{s,t}\|_{L_G^{\frac{q}{2}}} \leq C|t-s|^{2\beta}, \quad (30)$$

for some constant  $C$ . Then for all  $\alpha \in [0, \beta - \frac{1}{q})$ ,  $(X, \mathbb{X})$  has a quasi-surely modification, also denoted as  $(X, \mathbb{X})$ , and there exist  $K_\alpha \in L_G^q, \mathbb{K}_\alpha \in L_G^{\frac{q}{2}}$  such that for any  $s, t \in [0, T]$ , one has inequalities

$$|X_{s,t}| \leq K_\alpha |t-s|^\alpha, \quad |\mathbb{X}_{s,t}| \leq \mathbb{K}_\alpha |t-s|^{2\alpha}, \quad \hat{c} - q.s.. \quad (31)$$

Specially, if  $\beta - \frac{1}{q} > \frac{1}{3}$ , then quasi-surely  $\mathbf{X} = (X, \mathbb{X}) \in \mathcal{C}^\alpha([0, T], \mathbb{R})$ , for any  $\alpha \in (\frac{1}{3}, \beta - \frac{1}{q})$ .

*Proof.* Let  $T=1$ , and define dyadic partition as  $D_n = \{i2^{-n}, i = 0 \dots 2^n\}$ . Set

$$K_n = \sup_{t \in D_n} |X_{t, t+2^{-n}}|, \quad \mathbb{K}_n = \sup_{t \in D_n} |\mathbb{X}_{t, t+2^{-n}}|.$$

Note that since  $D_n$  are finite sets,  $K_n, \mathbb{K}_n$  belong to  $L_G^q$  and  $L_G^{\frac{q}{2}}$  respectively. Furthermore, one has bounds

$$\begin{aligned} \hat{\mathbb{E}}(K_n^q) &\leq \sum_{D_n} \hat{\mathbb{E}}|X_{t, t+2^{-n}}|^q \leq C^q \left(\frac{1}{2^n}\right)^{\beta q - 1} \\ \hat{\mathbb{E}}(\mathbb{K}_n^{\frac{q}{2}}) &\leq \sum_{D_n} \hat{\mathbb{E}}|\mathbb{X}_{t, t+2^{-n}}|^{\frac{q}{2}} \leq C^{\frac{q}{2}} \left(\frac{1}{2^n}\right)^{\beta q - 1} \end{aligned}$$

For any  $s, t \in \bigcup_n D_n$ , there exists  $m$  such that  $2^{-m-1} < t-s \leq 2^{-m}$ , and a partition,  $s = \tau_0 < \tau_1 < \dots < \tau_N = t$ , with  $(\tau_i, \tau_{i+1}) \in D_k$ , for some  $k \geq m+1$ . Also, we can choose such a partition that for any  $k \geq m+1$ , at most two such intervals are taken from  $D_k$ .

Then one obtains

$$|X_{s,t}| \leq \max_{0 \leq i \leq N} |X_{s, \tau_i}| \leq \sum_{i=0}^{N-1} |X_{\tau_i, \tau_{i+1}}| \leq 2 \sum_{n \geq m+1} K_n.$$

It follows that

$$\frac{|X_{s,t}|}{|t-s|^\alpha} \leq 2 \sum_{n \geq m+1} (2^n)^\alpha K_n \leq K_\alpha,$$

where  $K_\alpha := 2 \sum_{n \geq 0} 2^{n\alpha} K_n$ . We can easily check that  $K_\alpha \in L_G^q$ , since  $K_n \in L_G^q$ . For the second order part  $\mathbb{X}$ , Chen's identity plays a vital role in the calculation to cancel redundant terms implicitly. Indeed, it guarantees the first identity in

the following inequalities,

$$\begin{aligned}
|\mathbb{X}_{s,t}| &= \left| \sum_{i=0}^{N-1} (\mathbb{X}_{\tau_i, \tau_{i+1}} + X_{s, \tau_i} X_{\tau_i, \tau_{i+1}}) \right| \\
&\leq 2 \sum_{m+1} \mathbb{K}_n + \max_{0 \leq i \leq N} |X_{s, \tau_{i+1}}| \sum_{j=0}^{N-1} |X_{\tau_j, \tau_{j+1}}| \\
&\leq 2 \sum_{n \geq 1} \mathbb{K}_n + (2 \sum_{n \geq m+1} K_n)^2,
\end{aligned}$$

the last term of which can be checked to belong to  $L_G^{\frac{q}{2}}$  by similar argument.  $\square$

**Corollary 3.10.** (*Kolmogorov criterion for rough path distance in G-framework*)

Suppose  $\mathbf{X}, \tilde{\mathbf{X}}$  satisfy the moment condition as the above theorem. Let  $\Delta X = \tilde{X} - X$  and  $\Delta \mathbb{X} = \tilde{\mathbb{X}} - \mathbb{X}$  and assume that for some  $\varepsilon > 0$ , one has bounds,

$$\|\Delta X_{s,t}\|_{L_G^q} \leq C\varepsilon|t-s|^\beta, \|\Delta X_{s,t}\|_{L_G^{\frac{q}{2}}} \leq C\varepsilon|t-s|^{2\beta}.$$

Then there exists a constant  $M$ , depending on  $C, \alpha, \beta, q$ , such that

$$\|\|\Delta X\|_\alpha\|_{\mathbb{L}^q} \leq M\varepsilon, \|\|\Delta \mathbb{X}\|_{2\alpha}\|_{\mathbb{L}^{\frac{q}{2}}} \leq M\varepsilon.$$

In particular, if  $\beta - \frac{1}{q} > \frac{1}{3}$ , then for any  $\alpha \in (\frac{1}{3}, \beta - \frac{1}{q})$ ,  $\|\|\mathbf{X} - \tilde{\mathbf{X}}\|_{\mathcal{C}^\alpha}\|_{\mathbb{L}^q} \leq M\varepsilon$ .

*Proof.* By the same argument as the above theorem, there exists  $\Delta K_\alpha \in L_G^q$ , such that

$$\frac{|\Delta X_{s,t}|}{|t-s|^\alpha} \leq \varepsilon \Delta K_\alpha,$$

so it comes to the first inequality.

For the second inequality, note that Chen's identity fails for  $(\Delta X, \Delta \mathbb{X})$ . However, one still has the following estimate,

$$\begin{aligned}
|\Delta \mathbb{X}_{s,t}| &\leq \sum_{i=0}^{N-1} |\Delta \mathbb{X}_{\tau_i, \tau_{i+1}}| + \sum_{i=0}^{N-1} |X_{s, \tau_i} X_{\tau_i, \tau_{i+1}} - \tilde{X}_{s, \tau_i} \tilde{X}_{\tau_i, \tau_{i+1}}| \\
&\leq |\Delta \mathbb{K}_\alpha| |t-s|^{2\alpha} \varepsilon + \Delta K_\alpha K_\alpha \varepsilon |t-s|^{2\alpha} + \Delta K_\alpha \tilde{K}_\alpha \varepsilon |t-s|^{2\alpha},
\end{aligned}$$

with symbols adapted from the above proof. Then the bound for  $\|\|\Delta \mathbb{X}\|_{2\alpha}\|_{\mathbb{L}^{\frac{q}{2}}}$  follows.  $\square$

Denote  $\mathbf{B}^{strat} = (B, \mathbb{B}^{strat})$ , where  $\mathbb{B}_{s,t}^{strat} = \int_s^t B_{s,r} \circ dB_r$ . The lift of  $\mathbf{B}^{strat}$  as rough paths was done in [17] by applying the Kolmogorov's criterion. Now we will calculate the quasi-surely convergence rate for  $\mathbf{B}^{(n)}$  to  $\mathbf{B}^{strat}$ . Compared with the classical case, where Gaussian techniques could help a lot in the calculation of the second order paths, we do a direct calculation in the followings.

**Lemma 3.11.** Fix  $\alpha \in (\frac{1}{3}, \frac{1}{2})$ . Suppose  $T = 1$ ,  $B^{(n)}$  be the piecewise linearization as above, and  $\mathbf{B}^{(n)}$  be the natural enhancement of  $B^{(n)}$ . Then for any  $\theta < \frac{1}{2} - \alpha$  and  $q \geq 2$ ,  $\mathbf{B}^{(n)}$  converges to  $\mathbf{B}^{strat}$  under  $\alpha$ -Hölder rough norm in  $\mathbb{L}^q$  sense. Furthermore, one has the following inequalities,

$$\|\mathbf{B}^{strat} - \mathbf{B}^{(n)}\|_{\mathcal{C}^\alpha \mathbb{L}^q} \leq K \left(\frac{1}{n}\right)^\theta, \quad (32)$$

$$\varrho_\alpha(\mathbf{B}^{strat}, \mathbf{B}^{(n)}) \leq M \left(\frac{1}{n}\right)^\theta, \quad \hat{c} - q.s., \quad (33)$$

where  $K$  depends on  $\bar{\sigma}$  and  $M$  depends on  $\bar{\sigma}$  and the path  $\omega$ .

*Proof.* To show inequality (32), according to Corollary 3.10, by taking  $\beta = \frac{1}{2} - \theta$ , one only needs to show

$$\|\Delta B_{s,t}^{(n)}\|_{\mathbb{L}^q} \leq K \left(\frac{1}{n}\right)^\theta (t-s)^{\frac{1}{2}-\theta}, \quad (34)$$

$$\|\Delta \mathbb{B}_{s,t}^{(n)}\|_{\mathbb{L}^q} \leq K \left(\frac{1}{n}\right)^\theta (t-s)^{1-2\theta}, \quad (35)$$

for any  $q \geq 2$ , where  $\Delta B^{(n)} = B - B^{(n)}$  and  $\Delta \mathbb{B} = \mathbb{B} - \mathbb{B}^{(n)}$ .

Firstly, we prove (34).

Case1. If  $t_i^{(n)} \leq s < t \leq t_{i+1}^{(n)}$ , for some  $0 \leq i \leq n-1$ , one has

$$\Delta B_{s,t}^{(n)} = B_{s,t} + \frac{t-s}{t_{i+1}^{(n)} - t_i^{(n)}} B_{t_i^{(n)}, t_{i+1}^{(n)}}.$$

By basic inequality and Lemma 2.12, it follows that

$$\begin{aligned} \hat{\mathbb{E}}|\Delta B_{s,t}^{(n)}|^q &\leq 2^q [\hat{\mathbb{E}}|B_{s,t}|^q + \left(\frac{t-s}{t_{i+1}^{(n)} - t_i^{(n)}}\right)^q \hat{\mathbb{E}}|B_{t_i^{(n)}, t_{i+1}^{(n)}}|^q] \\ &\leq C_q \bar{\sigma}^q [|t-s|^{\frac{q}{2}} + |t-s|^q |t_{i+1}^{(n)} - t_i^{(n)}|^{-\frac{q}{2}}] \\ &\leq C_q \bar{\sigma}^q |t-s|^{\frac{q}{2}} \leq C_q \bar{\sigma}^q |t-s|^{\frac{q}{2}-q\theta} \left(\frac{1}{n}\right)^{q\theta}, \end{aligned}$$

where  $C_q$  depends only on  $q$ .

Case2. Suppose  $\tau_1 \leq s < \tau_2 < \dots < \tau_k < t \leq \tau_{k+1}$ , where  $\{\tau_i\}_{i=1}^{k+1} \subseteq \{t_i^{(n)}\}_{i=0}^n$  (we omit  $(n)$  in  $\tau_i$ ). Note that

$$\Delta B_{s,t}^{(n)} = \Delta B_{s,\tau_2}^{(n)} + \Delta B_{\tau_k,t}^{(n)},$$

and one could simply get the desired result by Case1.

Secondly, for (35), suppose  $t_i^{(n)} \leq \tau_0 := s < \tau_1 < \dots < \tau_{k-1} < t =: \tau_k \leq t_j^{(n)}$ , for some  $0 \leq i < j \leq n$  and  $k \geq 2$  (it's simple to check that the following is trivial if  $k = 1$ ), where  $\{\tau_i\}_{i=1}^{k-1} \subseteq \{t_i^{(n)}\}_{i=0}^n$ . According to Chen's identity (i.e.

equation (1)), one has the following identities,

$$\begin{aligned}\mathbb{B}_{s,t}^{(n)} &= \sum_{i=0}^{k-1} \mathbb{B}_{\tau_i, \tau_{i+1}}^{(n)} + \sum_{0 \leq i < j \leq k-1} B_{\tau_i, \tau_{i+1}}^{(n)} B_{\tau_j, \tau_{j+1}}^{(n)}, \\ \mathbb{B}_{s,t}^{strat} &= \sum_{i=0}^{k-1} \mathbb{B}_{\tau_i, \tau_{i+1}}^{strat} + \sum_{0 \leq i < j \leq k-1} B_{\tau_i, \tau_{i+1}} B_{\tau_j, \tau_{j+1}}.\end{aligned}$$

Then we get

$$\Delta \mathbb{B}_{s,t}^{(n)} = \sum_{i=0}^{k-1} \Delta \mathbb{B}_{\tau_i, \tau_{i+1}}^{(n)} + \sum_{0 \leq i < j \leq k-1} (B_{\tau_i, \tau_{i+1}} B_{\tau_j, \tau_{j+1}} - B_{\tau_i, \tau_{i+1}}^{(n)} B_{\tau_j, \tau_{j+1}}^{(n)}). \quad (36)$$

Note that by definition of  $\mathbb{B}^{(n)}$ , one has

$$\mathbb{B}_{\tau_i, \tau_{i+1}}^{(n)} = \int_{\tau_i}^{\tau_{i+1}} B_{\tau_i, r}^{(n)} dB_r^{(n)} = \frac{B_{\tau_i, \tau_{i+1}}^2}{2}, \quad i = 1, \dots, k-2,$$

while according to the definition of G-Stratonovich integral,

$$\mathbb{B}_{\tau_i, \tau_{i+1}}^{strat} = \int_{\tau_i}^{\tau_{i+1}} B_{\tau_i, r} \circ dB_r = \frac{B_{\tau_i, \tau_{i+1}}^2}{2}, \quad i = 0, \dots, k-1.$$

Then the first sum in (36) is actually

$$\sum_{i=0}^{k-1} \Delta \mathbb{B}_{\tau_i, \tau_{i+1}}^{(n)} = \mathbb{B}_{s, \tau_1}^{strat} - \mathbb{B}_{s, \tau_1}^{(n)} + \mathbb{B}_{\tau_{k-1}, t}^{strat} - \mathbb{B}_{\tau_{k-1}, t}^{(n)}. \quad (37)$$

According to Theorem 2.13, and Lemma 2.12, one has the following inequalities,

$$\hat{\mathbb{E}} |\mathbb{B}_{s, \tau_1}|^q \leq C_q \bar{\sigma}^q |\tau_1 - s|^q \leq C_q \bar{\sigma}^q \left(\frac{1}{n}\right)^q, \quad (38)$$

$$\hat{\mathbb{E}} |\mathbb{B}_{s, \tau_1}^{(n)}|^q \leq C_q \bar{\sigma}^{2q} |\tau_1 - s|^q \leq C_q \bar{\sigma}^{2q} \left(\frac{1}{n}\right)^q, \quad (39)$$

and similar results for  $\mathbb{B}_{\tau_{k-1}, t}^{strat}$  and  $\mathbb{B}_{\tau_{k-1}, t}^{(n)}$ .

As for the second sum in (36), note that  $B_{\tau_i}^{(n)} = B_{\tau_i}$ ,  $i = 1, \dots, k-1$ , and one

obtains

$$\begin{aligned}
& \sum_{0 \leq i < j \leq k-1} (B_{\tau_i, \tau_{i+1}} B_{\tau_j, \tau_{j+1}} - B_{\tau_i, \tau_{i+1}}^{(n)} B_{\tau_j, \tau_{j+1}}^{(n)}) \\
&= \sum_{1 \leq j \leq k-1} B_{s, \tau_1} B_{\tau_j, \tau_{j+1}} + \sum_{1 \leq i \leq k-2} B_{\tau_i, \tau_{i+1}} B_{\tau_{k-1}, \tau_k} \\
&- \sum_{1 \leq j \leq k-1} B_{s, \tau_1}^{(n)} B_{\tau_j, \tau_{j+1}}^{(n)} - \sum_{1 \leq i \leq k-2} B_{\tau_i, \tau_{i+1}}^{(n)} B_{\tau_{k-1}, \tau_k}^{(n)} \\
&= B_{s, \tau_1} B_{\tau_1, \tau_{k-1}} + B_{s, \tau_1} B_{\tau_{k-1}, t} + B_{\tau_1, \tau_{k-1}} B_{\tau_{k-1}, t} \\
&- B_{s, \tau_1}^{(n)} B_{\tau_1, \tau_{k-1}}^{(n)} - B_{s, \tau_1}^{(n)} B_{\tau_{k-1}, t}^{(n)} - B_{\tau_1, \tau_{k-1}}^{(n)} B_{\tau_{k-1}, t}^{(n)} \\
&= (B_{s, \tau_1} - B_{s, \tau_1}^{(n)}) B_{\tau_1, \tau_{k-1}} + B_{\tau_1, \tau_{k-1}} (B_{\tau_{k-1}, t} - B_{\tau_{k-1}, t}^{(n)}) + B_{s, \tau_1} B_{\tau_{k-1}, t} - B_{s, \tau_1}^{(n)} B_{\tau_{k-1}, t}^{(n)} \\
&= (B_s^{(n)} - B_s) B_{\tau_1, \tau_{k-1}} + B_{\tau_1, \tau_{k-1}} (B_t - B_t^{(n)}) + B_{s, \tau_1} B_{\tau_{k-1}, t} - B_{s, \tau_1}^{(n)} B_{\tau_{k-1}, t}^{(n)}. \tag{40}
\end{aligned}$$

Indeed, by independence, one can check that

$$\begin{aligned}
\hat{\mathbb{E}}|(B_s^{(n)} - B_s) B_{\tau_1, \tau_{k-1}}|^q &\leq \bar{\sigma}^q |\tau_{k-1} - \tau_1|^{\frac{q}{2}} \hat{\mathbb{E}}|(B_s^{(n)} - B_s)|^q \leq C_q \bar{\sigma}^{2q} \left(\frac{1}{n}\right)^{\frac{q}{2}} |t - s|^{\frac{q}{2}} \\
\hat{\mathbb{E}}|B_{s, \tau_1} B_{\tau_{k-1}, t}|^q &\leq \bar{\sigma}^{2q} \left(\frac{1}{n}\right)^q,
\end{aligned}$$

and similar results for  $B_{\tau_1, \tau_{k-1}} (B_t - B_t^{(n)})$  and  $B_{s, \tau_1} B_{\tau_{k-1}, t}^{(n)}$ .

By (37), (40), and applying Minkowski inequality, one obtains

$$\begin{aligned}
\|\Delta \mathbb{B}_{s, t}^{(n)}\|_{\mathbb{L}^q} &\leq C_q (\bar{\sigma}^2 + \bar{\sigma}) \left(\frac{1}{n}\right)^{\frac{1}{2}} |t - s|^{\frac{1}{2}} \\
&\leq C_q (\bar{\sigma}^2 + \bar{\sigma}) \left(\frac{1}{n}\right)^{\theta} |t - s|^{1-2\theta}.
\end{aligned}$$

By the randomness of  $q$ , one can get the following inequality from 3.10,

$$\|\varrho_\alpha(\mathbf{B}^{strat}, \mathbf{B}^{(n)})\|_{\mathbb{L}^q} \leq K \left(\frac{1}{n}\right)^\theta.$$

To prove (33), for any  $\theta < \frac{1}{2} - \alpha$ , one may choose  $q > 2$  and  $\theta < \theta' < \frac{1}{2} - \alpha$ , such that  $q(\theta' - \theta) > 2$ . It is clear that

$$\begin{aligned}
\hat{c} \left( \bigcap_{M=1} \bigcap_{n=1} \bigcup_{m \geq n} \{ \varrho_\alpha(\mathbf{B}^{strat}, \mathbf{B}^{(m)}) > M \frac{1}{m^\theta} \} \right) &\leq \hat{c} \left( \bigcap_{n=1} \bigcup_{m \geq n} \{ \varrho_\alpha(\mathbf{B}^{strat}, \mathbf{B}^{(m)}) > \frac{1}{m^\theta} \} \right) \\
&= \hat{c} \left( \limsup_m \{ \varrho_\alpha(\mathbf{B}^{strat}, \mathbf{B}^{(m)}) > \frac{1}{m^\theta} \} \right).
\end{aligned}$$

Note that

$$\hat{c}(\varrho_\alpha(\mathbf{B}^{strat}, \mathbf{B}^{(m)}) > \frac{1}{m^\theta}) \leq \frac{(\frac{1}{m^{\theta'}})^q}{(\frac{1}{m^\theta})^q} = \frac{1}{m^{q(\theta' - \theta)}}.$$

According to Borel-Cantelli argument in G-framework (see Lemma 5 in [1]), one obtains

$$\hat{c}(\bigcap_{M=1} \bigcap_{n=1} \bigcup_{m \geq n} \{\varrho_\alpha(\mathbf{B}^{strat}, \mathbf{B}^{(m)}) > M \frac{1}{n^\theta}\}) = 0,$$

which implies

$$\varrho_\alpha(\mathbf{B}^{strat}, \mathbf{B}^{(m)}) \leq M(\omega) \frac{1}{n^\theta}, \quad \hat{c} - q.s..$$

□

**Remark 3.12.** According to the above argument, it is possible to get sharper estimate by considering G-Brownian motion as rough paths with lower regularity (see [5] for the classical case).

**Remark 3.13.** One should note that the above calculation only happens in one-dimensional case. For the multiple-dimension case, the main problem is probably that components of G-Brownian motion, unlike the classical case, are not usually independent. Actually, mutual independency means linearity in G-framework. See [7] for the independency in G-framework.

Here is the main result of this section.

**Theorem 3.14.** Suppose  $f, g, h \in \mathcal{C}_b^3$ , and  $Y^{(n)}$  defined as in (18). Also, suppose  $X$  solves the G-Stratonovich SDE of (16) and  $Y$  solves the following RDE driven by G-Stratonovich rough paths,

$$dY_t = f(Y_t)d\mathbf{B}^{strat} + g(Y_t)d\langle B \rangle_t + h(Y_t)dt, \quad (41)$$

with initial condition as scalar  $x_0$ . Then for any  $\theta < \frac{1}{2} - \alpha$ , one has the following inequality,

$$\|Y - Y^{(n)}\|_\alpha \leq M(\omega) \frac{1}{n^\theta}, \quad \hat{c} - q.s..$$

In particular,  $X = Y$ ,  $\hat{c} - q.s.$ , and

$$\|X - Y^{(n)}\|_\alpha \leq M(\omega) \frac{1}{n^\theta}, \quad \hat{c} - q.s..$$

*Proof.* Apply Theorem 2.5 and Lemma 3.11, and one obtains the first inequality. Also, note that the particular part follows from Theorem 3.7 and the fact that  $X, Y$  are quasi-surely continuous with respect to  $t$ . □

The following corollary implies the continuity of solutions of RDEs driven by lifted G-Brownian motion with respect to uniform norm on the canonical space.

**Corollary 3.15.**  $Y$  is RDE solutions defined as the above theorem, then for any  $t < T$ ,  $Y_t$  has a quasi-continuous version.

*Proof.* It is clear by  $Y_t \in L_G(\Omega_t)$  and the representation of  $L_G(\Omega_t)$ . □

## References

- [1] Denis, L., Hu, M. and Peng, S.(2011) *Function spaces and capacity related to a sublinear expectation: application to G-Brownian motion pathes*. Potential Anal., 34: 139-161.
- [2] Deya, A., Neuenkirch, A. and Tindel, S.(2012) *A Milstein-type scheme without Lévy area terms for SDEs driven by fractional Brownian motion*. Ann. Inst. Henri Poincaré Probab. Stat. 48, no.2,518-550.
- [3] Friz, P. and Hairer, M.(2014) *A course on rough paths*. E-print available at <http://www.hairer.org/notes/RoughPaths.pdf>.
- [4] Friz, P. and Victoir, N.(2010) *Multidimensional stochastic processes as rough paths*. Vol.120 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge.
- [5] Friz, P and Riedel, S.(2014) *Convergence rates for the full Gaussian rough paths*. Ann. Inst. Henri Poincaré Probab. Stat. 50, no.1,154-194.
- [6] Hu, Y. and Nualart, D.(2009) *Rough path analysis via fractional calculus*. Trans. Amer. Math. Soc.361, no.5,2689-2718.
- [7] Hu, M. and Li, X.(2014) *Independence under the G-expectation framework*. J. Theor. Probab. 27:1011-1020.
- [8] Gubinelli, M.(2004) *Controlling rough paths*. Journal of Functional Analysis, 216(2004)86-140.
- [9] Gubinelli, M.(2010) *Ramification of rough paths*. J. Differential Equations 248, no.4,(2010),693-721.
- [10] Geng, X. Qian, Z and Yang, D.(2014) *G-Brownian motion as rough paths and differential equations driven by G-Brownian motion*. Séminaire de Probabilités XLVI, Lecture Notes in Mathematics 2123, DOI 10.1007/978-3-319-11970-0\_6. Springer, New York, London.
- [11] Ikeda, N and Watanabe, S.(1989) *Stochastic differential equations and diffusion processes*. North-Holland Publishing Co., Amsterdam, second ed..
- [12] Lyons, T.(1998) *Differential equations driven by rough signals*. Rev.Mat.Iberoamericana14,no.2,215-310.
- [13] Lyons, T. and Qian, Z.(2002) *System control and rough paths*. Oxford University Press, Oxford, et al.
- [14] Peng, S.(2007) *G-expectation, G-Brownian motion and related stochastic calculus of Itô type*. Stochastic Analysis and Applications, 541-567, Abel Symp., 2, Springer, Berlin.

- [15] Peng, S.(2008) *Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation*. Stochastic Processes and their Applications, 118(12): 2223-2253.
- [16] Peng, S.(2010) *Nonlinear expectations and stochastic calculus under uncertainty*. Available at arXiv:1002.4546v1 [math.PR].
- [17] Peng, S and Zhang, H.(2015) *G-stochastic calculus viewed through rough paths and the Norris lemma in G-framework*. ArXiv e-prints, arXiv:1510.01851.
- [18] Wong, E and Zakai, M.(1965) *On the relationship between ordinary and stochastic differential equations*. Int. J. Eng. Sci., 3, 213-29.
- [19] Wong, E and Zakai, M.(1965) *On the convergence of ordinary integrals to stochastic integrals*. Ann. Math. Stat., 36, 1560-4.